# Cryptography – The Secret Language of Spies

By Dan Veeneman

*"23187... 46982... 69335... 15948..."*

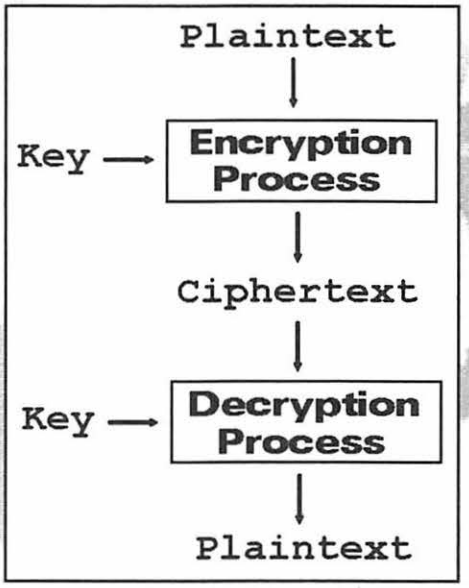
If you've listened to shortwave for any length of time you've probably run across them. A mysterious voice announcing a long series of numbers, perhaps destined for some unknown spy in a foreign land. Although shrouded in secrecy, these numbers are assumed to be some kind of encrypted message, making sense only after being decoded by some clandestine listener.

Cryptography is the science of secure communication in the presence of adversaries. In any encryption system, the message to be sent, called the *plaintext*, is encrypted in some way to produce *ciphertext*, which is then transmitted. Encryption systems typically require a secret piece of information, called a *key*, in order to scramble the message. On the receiving end, the key is used to decrypt the ciphertext and produce the original plaintext. Should the ciphertext be intercepted, an eavesdropper would be unable to decrypt it without the proper key, even if the encryption method was known.

Despite the lack of concrete details about these mysterious numbers stations, over the years it has been assumed that the numbers are really messages encrypted using a *one-time pad*.

## One-Time Pad

The only provably secure cryptographic system was created by Gilbert Vernam in 1917, who originally developed it to protect electronically transmitted messages. His invention automatically enciphered individual characters as they were entered into a teletypewriter as well as automatically deciphering them at the receiving end. This kind of character-by-character encryption is called a *stream cipher*, and Vernam's particular method is unbreakable if used correctly. His invention is now referred to as the one-time pad cryptosystem, and has been extended to both manual and computerized operations.




The key in a one-time pad system is a series of random numbers. Vernam's invention used long strips of paper tape, but in the manual version of the system numbers are printed on small pads of paper that are easily concealed. Each sheet of paper in the pad is used only once, then discarded and permanently destroyed. (It has been reported that the Central Intelligence Agency has a special type of paper that turns into chewing gum when it contacts saliva.) There are two copies of the one time pad: the person wishing to send a message and the intended recipient both have identical pads.

## An Example

With all that in mind, let's take a look at a contrived example.

Spy headquarters is preparing to send an operative out into the field. During training the operative learned that most cryptographic systems involve mathematical operations, where it's much easier to deal with numbers than letters. Spy HQ has established a simple coding scheme where each letter of the alphabet is assigned a number. The letter A has a value of 1, the letter B has a value of 2, and so on up to Z, which has a value of 26 (see Table 1).

**Table 1: Letter-number pairings.**

| Letter | Value |
|---|---|
| A | 1 |
| B | 2 |
| C | 3 |
| ... | ... |
| Z | 26 |

In addition to a false passport and some local currency, the operative is issued a small pad of paper with a series of random letters printed on each page. This one-time pad contains the keys for the operative to decrypt messages that will be sent in the future. Spy HQ keeps an identical copy of the pad. Some time passes and the operative becomes active in the target country, occasionally listening to shortwave radio for any messages that may be destined for him.

At some point Spy HQ wishes to send a message instructing him to purchase a particular magazine. Using the letter-to-number conversion table, Spy HQ converts the phrase MONITORING TIMES into the numbers 13 15 14 09 20 15 18 09 14 07 20 09 13 05 19 (skipping the space). This is the plaintext message.

To encrypt the message Spy HQ will use a page from the one-time pad. For this example we'll say the next page of the pad has the key SECRET LISTENERS written on it, although

in practice the key would be a random string of numbers. Again ignoring spaces, SECRET LISTENERS is encoded as 19 05 03 18 05 20 12 09 19 20 05 14 05 18 19.

The simplest encryption method is to add the plaintext number to the corresponding key number for each letter in the message (see Figure 1).

```
13 15 14 09 20 15 18 09 14 07 20 09 13 05  19
MONITORING TIMES
19 05 03 18 05 20 12 09 19 20 05 14 05 18  19
SECRET LISTENERS
-- -- -- -- -- -- -- -- -- -- -- -- -- --  --
32 20 17 27 25 35 30 18 33 27 25 23 18 23  38
```

**FIGURE 1:** *Adding each plaintext number to the corresponding key number.*

Since the letter encoding scheme only goes up to 26, Spy HQ "wraps around" the sums that exceed 26 by subtracting 26 and using the difference (see Figure 2). This operation is called taking the modulus and is done in many situations involving counting. (For instance, there are twelve hours on the face of a normal clock, and after reaching 12 we "wrap around" and continue counting at 1. This is called taking the hour "modulo 12.")

```
32 20 17 27 25 35 30 18 33 27 25 23 18 23  38
-26 -26 -26-26 -26-26 -- -- -- -- --  -26
06 20 17 01 25 09 04 18 07 01 25 23 18 23  12
```

**FIGURE 2:** *Taking the addition result modulo 26.*

So Spy HQ's encoded message is 06 20 17 01 25 09 04 18 07 01 25 23 18 23 12 (the nonsensical FTQAYIDRGAYWRWL), which is transmitted over the air at a time when the operative is scheduled to be listening.

On the receiving end the operative uses the identical key from his one-time pad to recover the message. Each key value is subtracted from the corresponding ciphertext letter to produce the plaintext (see Figure 3).

```
06 20 17 01 25 09 04 18 07 01 25 23 18 23  12
19 05 03 18 05 20 12 09 19 20 05 14 05 18  19
-13 15 14-17 20-11-08 09-12-19 20 09 13 05 -07
```

**FIGURE 3:** *Decoding the received message by subtraction.*

The values that fall below 1 are "wrapped around" to a positive value by adding 26 (Figure 4).

```
-13 15 14-17 20-11-08 09-12-19 20 09 13 05 -07
+26  +26 +26+26 +26+26  --  --  +26
13 15 14 09 20 15 18 09 14 07 20 09 13 05 12
```

**FIGURE 4:** *Taking the difference modulo 26.*

This gives the operative the original plaintext message of MONITORING TIMES.

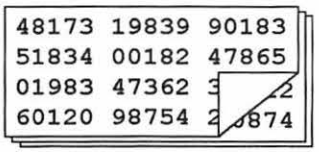
An eavesdropper listening to the same shortwave station would have just the numbers 06 20 17 01 25 09 04 18 07 01 25 23 18 23 12, spelling FTQAYIDRGAYWRWL. If Spy HQ and the operative are following all the rules to maintain security, the best the eavesdropper can do is try out every possible key. This method is called *brute force* decryption, but in a one-time pad system where the key is as long as the message, even that drastic step won't help. The eavesdropper, in fact, will end up with all possible solutions.

For example, decrypting FTQAYIDR GAYWRWL using the key NSNVVHL YTLKCJKM produces the result RACE CARS MONTHLY and the key RSWRJUCFMZECFRT will produce NATIONAL TATTLER, both of which make sense but neither of which are correct. Since any key in a secure one-time pad is equally likely, the eavesdropper has no hope of determining the true plaintext using brute force.

## Keeping it Secure

In order for a one-time pad system to remain secure, several rules must be followed:

**Example One-Time Pad**




First, the contents of the one-time pads must be kept secret. This is fairly obvious, but important. Anyone discovering the pads, if they didn't shoot the owner on sight for being a spy, could make a copy of the pad and decrypt the messages themselves. Intelligence and law enforcement agencies routinely perform "black bag jobs" to covertly break in and steal keys used by embassies, corporations, and even private individuals. This is probably the most damaging situation for the one-time pad user, since the only thing worse than no security is a false sense of security.

Second, the pads must not be reused. Two-time pads are not mathematically secure, and in fact the reuse of pads allowed the National Security Agency (NSA) to break thousands of Soviet KGB and GRU messages in the 1940's.

Third, the contents of the pads must be unpredictable. Regardless of whether the spy keeps the pads a secret and doesn't reuse them, if the numbers are somehow predictable it would be the equivalent of the adversary having their own copy of the pad. Numerous modern software programs that offer "one-time pads" have this weakness — their method of creating "random" numbers is often pre-

dictable, allowing a snoop to reproduce the key and decode messages.

### Randomness

Do these numbers seem random and unpredictable: 212, 198, 216, 32, 175, 100? If you think so, you don't commute on Interstate 95 north of Washington, D.C. These are the Maryland route numbers, in order, of the exits off northbound I-95 after leaving the D.C. beltway. They may look random, but they are very predictable once you know the pattern.

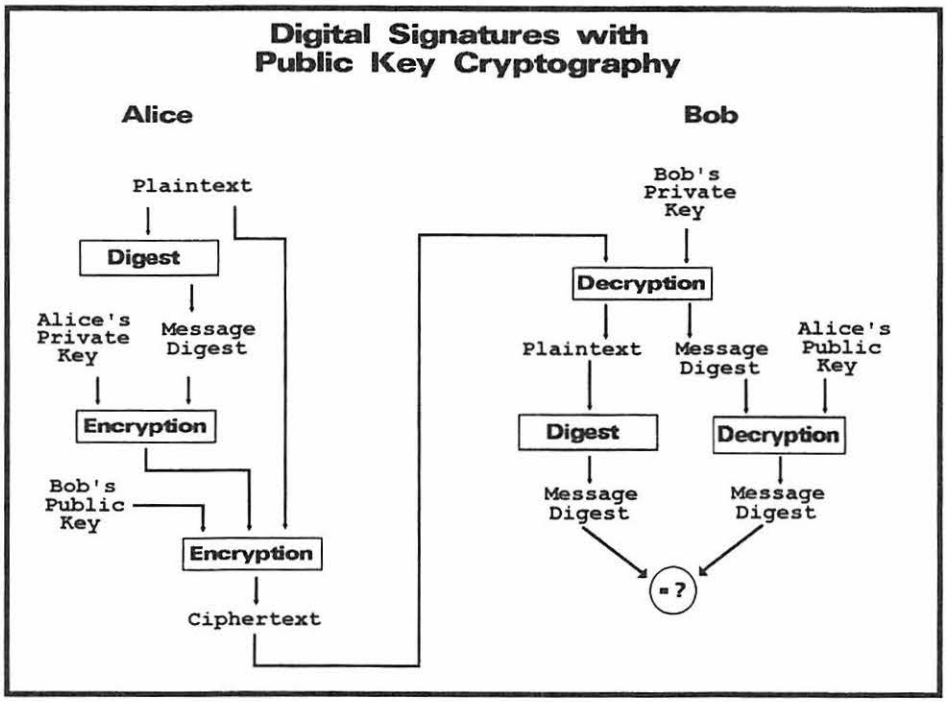
It's surprisingly difficult to generate really unpredictable numbers. For those of you who may be planning to use the random number function available in most programming languages for cryptography, please think again. Just like the Maryland route numbers, the output from these functions look random but are predictable and repeatable. In fact, a couple of years ago a serious weakness in Netscape's web browser was revealed after researchers discovered the random number generator used in the browser was very weak. This weakness would have allowed eavesdroppers to easily decode protected web connections, including the details of any financial transactions.

### Key Length

In a secure one-time pad system the key must be as long as the message. If you want to send a message containing two thousand characters, you need a key that is also 2,000 characters long. If you want to encrypt a computer file that is two megabytes (2 million characters) with a one-time pad, you will need a key that is also two megabytes. If you want to protect a digital video signal that carries hundreds of thousands of characters each second, you probably won't use a one-time pad. Creating, delivering, and securing large keys is a very difficult problem, limiting true one-time pad systems to very specific uses, such as clandestine operatives in foreign countries.

For the rest of us, practical limits on the amount of key information we can handle necessitate the use of other encryption systems. In *block ciphers* the key length is much less than the length of the message, and is used in a different way. The plaintext message is broken up into small pieces called blocks, and each block is encrypted by the key. There are a wide variety of block ciphers that operate in different ways, but as a general rule the smaller the key the less secure any message encrypted with that key will be.

Keys in modern cryptographic systems are measured by the number of bits they contain. A bit is the smallest unit of information, either a 0 or a 1. The Data Encryption Standard (DES), a popular cipher, uses keys that are 56 bits long, allowing approximately 72 quadrillion (72,057,594,037,927,936) possible keys. As we'll see, this isn't nearly enough. Skipjack, a cipher invented inside the NSA and classified secret until very recently, uses 80 bit keys. The International Data Encryption Algorithm (IDEA) uses keys that are 128 bits long. Other ciphers of varying strength and complexity exist as well, each with their own key lengths. All other things being equal, one additional bit doubles the amount of work necessary to brute force a solution.




The United States government considers cryptography to be an implement of war, and closely controls the export of strong cryptography. Until recently, ciphers using more than 40 bits of key were usually denied export permits with the justification that such systems were too difficult for the government to crack and would therefore jeopardize national security. Even 56-bit DES has been denied export in many circumstances, despite its availability overseas. In addition, domestic law enforcement representatives have been wringing their hands in public, fretting that the widespread use of strong cryptography would render their wiretaps useless and frustrate their investigative efforts. In an effort to convince law makers of the dangers of strong cryptography, the Federal Bureau of Investigation (FBI) has testified before Congress numerous times about the impossibility of cracking encrypted messages protected by cipher systems such as DES.

### DES Cracking

The Data Encryption Standard (DES) is a Federal Information Processing Standard (FIPS) approved by the National Bureau of Standards (now the National Institute of Standards and Technology, NIST) in 1977. In the intervening decades it has been scrutinized by numerous experts in cryptography, none of whom have found a significant weakness in its fundamental design. What is weak is the inadequate size of the key.

As mentioned above, brute force cracking is the process of decrypting a message using every possible key to find the one that works. To motivate practical research, in 1997 a company named RSA Data Security offered a prize for cracking a message encrypted using DES. Hundreds of volunteers ran customized key search software in their spare time on available computers, exchanging results and passing key information over the Internet. The message was broken after five months of effort across several thousand computers. At the beginning of this year the prize was offered again with a different message, and an improved search method yielded the key in only 39 days, again broken by a group effort of thousands of computers coordinated over the Internet.

A third challenge was broken in July by the Electronic Frontier Foundation (EFF), who used a single custom-built machine to crack the code in less than three days. For well under $250,000 the EFF had built a "DES Cracker" from a personal computer and an array of

custom microchips that could break a DES-encrypted message in reasonable time without having to coordinate with anyone. They also showed that the government had been playing fast and loose with the truth in their testimony of how difficult it is to break DES messages.

## Public Key Systems

Existing encryption systems fall into two categories. Private key systems, sometimes called secret key systems, use the same key on the sending and the receiving ends. One-time pad and DES cipher methods are private key systems.

Public key systems, on the other hand, make use of a public/private key pair. The public key is published and available for anyone to see while the private key is kept secret by the owner. The public and private keys are mathematically related to each other, but for all practical purposes it is impossible for an adversary given the public key to determine the private key.

For an overview of the use of a public key system let me introduce some characters that have become standard in the cryptologic literature. The protagonists in our story are Alice and Bob, who want to communicate securely while an adversary, say Louis, wishes to intercept and read everything they send to each other.

To send a message in a public key system, Alice creates a message destined for Bob. She encrypts the message using Bob's public key and sends it to him. When Bob receives the message he uses his private key and decodes the message. Since no one else knows Bob's private key, no one else will be able to decode the message.

Although Alice is sure that Bob is the only one that can read the message, how can Bob be sure the message really came from Alice? Perhaps our snoop Louis has effected a *man-in-the-middle* attack and intercepted Alice's message, altering it or replacing it with another message of his own creation before sending it on to Bob.

## Digital Signatures

So far we've been talking about *confidentiality*, that is, keeping the contents of the message secret. Cryptography can also help us with *authentication*, that is, proving that a message really came from a particular person.

To foil Louis, Alice can take some additional steps that will prove to Bob she is really the author of the message he receives. Before sending the message, Alice generates what's called a *message digest*, a kind of digital fingerprint that identifies the contents of the message. This digest is based on the exact contents of the message, and no two messages will have the same digest. Alice then encrypts the digest with her private key and sends it along with the encrypted message to Bob. This is the equivalent of Alice signing the letter, and in fact this called a *digital signature*.

As before, Bob receives the encrypted message and decrypts it using his private key. He also decrypts the encrypted digest using Alice's public key. Bob then runs the plaintext message through a message digest function to produce a local fingerprint. If the digest decrypted with Alice's public key matches the digest he just computed, Bob can be sure that the message was not altered and was, in fact, sent by Alice and not an impostor.

All of these features are available in a program called *Pretty Good Privacy*, which was one of the first publicly-available programs to provide strong encryption. PGP uses public key cryptography to protect message confidentiality and assure authentication of the sender. (*MT's Computers & Radio also reviewed a data encryption program called Cyberlock in the August issue - ed.*)

## A Challenge

As a final challenge, see if you can decrypt the following one-time pad message. Spy HQ has sent you the message **BABPFQH WZGIRPFNWBCOS** and the next key on your one-time pad is **LISTENING AND LEARNING**.

What is Spy HQ trying to remind you? The answer will appear in my *PCS Front Line* column in the December issue of *Monitoring Times*. For those of you who can't wait, the answer will also be available on my website, which will also have a computer program to assist you. Details below.

## Further Reading

For more on the history of code making and code breaking, David Kahn's *The Codebreakers* is an authoritative account of the use and misuse of cryptography.

For those more technically inclined, Bruce Schneier's *Applied Cryptography* is one of the most popular books detailing modern cryptographic systems and how to implement them.

The American Cryptogram Association is a non-profit, volunteer organization dedicated to the creation and solution of cryptographic challenges of various sorts. More information on the ACA is available by writing to ACA Treasurer, 1118 Via Palo Alto, Aptos, CA, 95003.

All of this information and more is available on my website at **www.decode.com**.